Sources of bias in artificial intelligence that perpetuate healthcare disparities—A global review

Year: 2022 | Citations: 311 | Authors: L. Celi, J. Cellini, Marie Charpignon, E. Dee, Franck Dernoncourt

Abstract

Background While artificial intelligence (AI) offers possibilities of advanced clinical prediction and decision-making in healthcare, models trained on relatively homogeneous datasets, and populations poorly-representative of underlying diversity, limits generalisability and risks biased AI-based decisions. Here, we describe the landscape of AI in clinical medicine to delineate population and data-source disparities. Methods We performed a scoping review of clinical papers published in PubMed in 2019 using AI techniques. We assessed differences in dataset country source, clinical specialty, and author nationality, sex, and expertise. A manually tagged subsample of PubMed articles was used to train a model, leveraging transfer-learning techniques (building upon an existing BioBERT model) to predict eligibility for inclusion (original, human, clinical AI literature). Of all eligible articles, database country source and clinical specialty were manually labelled. A BioBERT-based model predicted first/last author expertise. Author nationality was determined using corresponding affiliated institution information using Entrez Direct. And first/last author sex was evaluated using the Gendarize.io API. Results Our search yielded 30,576 articles, of which 7,314 (23.9%) were eligible for further analysis. Most databases came from the US (40.8%) and China (13.7%). Radiology was the most represented clinical specialty (40.4%), followed by pathology (9.1%). Authors were primarily from either China (24.0%) or the US (18.4%). First and last authors were predominately data experts (i.e., statisticians) (59.6% and 53.9% respectively) rather than clinicians. And the majority of first/last authors were male (74.1%). Interpretation U.S. and Chinese datasets and authors were disproportionately overrepresented in clinical AI, and almost all of the top 10 databases and author nationalities were from high income countries (HICs). AI techniques were most commonly employed for image-rich specialties, and authors were predominantly male, with non-clinical backgrounds. Development of technological infrastructure in data-poor regions, and diligence in external validation and model re-calibration prior to clinical implementation in the short-term, are crucial in ensuring clinical AI is meaningful for broader populations, and to avoid perpetuating global health inequity.